Multimedia Systems Lecture – 15

By

Dr. Priyambada Subudhi

Assistant Professor

IIIT Sri City

3D Video and TV

- Three-dimensional (3D) pictures and movies have been in existence for decades.
- Increasingly, it is in movie theaters, broadcast TV (e.g., sporting events), personal computers, and various handheld devices.
- The main advantage of the 3D video is that it enables the experience of immersion— be there, and really Be there.
- We will see fundamentals of 3D vision or 3D percept, emphasizing stereo vision (or stereopsis) since most modern 3D video and 3D TV are based on stereoscopic vision.

Cues for 3D Percept

- The human vision system is capable of achieving a 3D percept by utilizing multiple cues.
- They are combined to produce optimal (or nearly optimal) depth estimates.
- When the multiple cues agree, this enhances the 3D percept. When they conflict with each other, the 3D percept can be hindered. Sometimes, illusions can arise.

Monocular Cues

- The monocular cues that do not necessarily involve both eyes include:
  - Shading—depth perception by shading and highlights
  - Perspective scaling—converging parallel lines with distance and at infinity
  - Relative size—distant objects appear smaller compared to known same-size objects not in distance
  - Texture gradient—the appearance of textures change when they recede in distance
  - Blur gradient—objects appear sharper at the distance where the eyes are focused, whereas nearer and farther objects are gradually blurred
  - Haze—due to light scattering by the atmosphere, objects at distance have lower contrast and lower color saturation
  - Occlusion—a far object occluded by nearer object(s)
  - Motion parallax—induced by object movement and head movement, such that nearer objects appear to move faster.
- Among the above monocular cues, it has been said that Occlusion and Motion parallax are more effective.

Binocular Cues

- The human vision system utilizes effective binocular vision, i.e., *stereo vision.*
- Our left and right eyes are separated by a small distance, on average approximately 2.5 inches, or 65mm. This is known as the *interocular distance*.
- As a result, the left and right eyes have slightly different views, i.e., images of objects are shifted horizontally.
- The amount of the shift, or *disparity*, is dependent on the object's distance from the eyes, i.e., its *depth*, thus providing the binocular cue for the 3D percept.
- The horizontal shift is also known as *horizontal parallax*.
- The fusion of the left and right images into single vision occurs in the brain, producing the 3D percept.

Stereo Camera Model

- We can design a simple (artificial) stereo camera system in which the left and right cameras are identical (same lens, same focal length, etc.); the cameras' optical axes are in parallel, pointing at the *Z*-direction, the scene depth.
- The cameras are placed at $(-b/2, 0, 0)$ and $(b/2, 0, 0)$ in the world coordinate system where *b* is camera separation, or the length of the *baseline* where *b* is camera separation, or the length of the *baseline*.
- Given a point $P(X, Y, Z)$ in the 3D space, and x_l and x_r being the *x*-coordinates of its projections on the left and right camera image planes, the following can be derived:

  $$d = f\, b/Z,$$

where *f* is the focal length, $d = x_l - x_r$ is the *disparity* or *horizontal parallax*.

- This suggests that disparity d is inversely proportional to the depth Z of the point P.
- Namely, objects near the cameras yield large disparity values, and far objects yield small disparity values. When the point is very far, approaching infinity, $d \rightarrow 0$.
- Moreover, objects at the same depth in the scene will have the same disparity d. This enables us to depict the 3D space with a stack of *depth planes*, or equivalently, *disparity planes*, which is handy in camera calibration, video processing and analysis.

3DMovie and TV Based on Stereo Vision

- **3D Movie Using Colored Glasses**
- **3D Movies Using Circularly Polarized Glasses**
- **3D TV with Shutter Glasses**